Protonated α- N -Acetyl Galactose Glycopeptide Dissociation Chemistry

Jordan Rabus, Shanshan Guan, Lauren Schultz, Maha Abutokaikah, Philippe Maître, Benjamin Bythell

Protonated α-GalNAc-Glycopeptide Dissociation Chemistry

Jordan M. Rabus,[†,‖,Δ] Shanshan Guan,[†,‖,Δ] Lauren M. Schultz,[†] Maha T. Abutokaikah,[‖] Philippe Maître,[‡] and Benjamin J. Bythell[†,‖*]

[†]Department of Chemistry and Biochemistry, Ohio University, 307 Chemistry Building, OH 45701, USA.

[‖]Department of Chemistry and Biochemistry, University of Missouri-St. Louis, 1 University Blvd, St. Louis, MO 63121, USA.

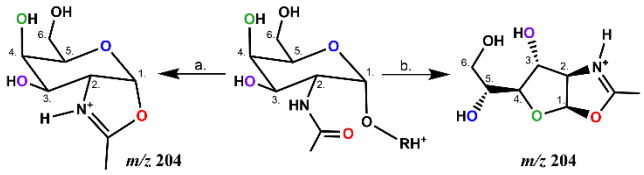
[‡]Université Paris-Saclay, CNRS, Institut de Chimie Physique, 91405, Orsay, France.

ABSTRACT: We recently provided mass spectrometric, H/D labeling, and computational evidence of pyranose to furanose N-Acetylated ion isomerization reactions occurring prior to glycosidic bond cleavage in both O- and N-linked glycosylated amino acid model systems (*Phys. Chem. Chem. Phys.*, **2021**, 23, 23256-23266). These reactions occurred irrespective of glycosidic linkage stereochemistry (α/β) and N-Acetylated hexose structure (GlcNAc/GalNAc). In the present article we test the generality of the preceding findings by examining Threonyl α-GalNAc glycosylated peptides. We utilize computational chemistry to compare the various dissociation and isomerization pathways accessible with collisional activation. We then interrogate the structure(s) of the resulting charged glycan and peptide fragments with infrared "action" spectroscopy. Isomerization of the original pyranose, protonated glycopeptide, $[AT(GalNAc)A+H]^+$, is predicted to be facile compared to direct dissociation, as is glycosidic bond cleavage of the newly formed furanose form. i.e., furanose oxazolinium ion structures are predicted to predominate. IR action spectra for the *m/z* 204, $C_8H_{14}N_1O_5^+$, glycan fragment population support this prediction. The IR action spectra of the complementary *m/z* 262, peptide fragment are assigned as a mixture of the lowest energy structures of $[ATA+H]^+$ consistent with the literature. If general, the change to a furanose m/z 204 product ion structure fundamentally alters the ion population available for MS^3 dissociation and glycopeptide sequence identification.

Peptide and protein glycosylation is widespread in biological systems. Glycosylation is involved in the pathogenesis of diseases, with variations of glycan at each particular site thought to provide further biomarkers of disease[1–9] and immune response.[10,11] The spike, envelope, and membrane structural proteins of SARS-CoV-2 and other coronaviruses are glycosylated, affecting their attachment to host, entry, replication, and infection.[12,13] These also affect the viruses ability to mask the proteins from the body's immune system.[13] Consequently accurate profiling of site-specific glycan modifications has great importance.[14–17]

Despite numerous studies/approaches involving these analytes' there are still many challenges to glycoprotein sequencing.[9,14] Current large-scale sequencing and identification of glycopeptides is contingent on algorithms almost wholly reliant on m/z differences between the fragment ion series.[18–20] Thus our ability to differentiate a hexose (Hex) from a N-Acetylated-hexose (HexNAc) residue or fragment, far outstrips our ability to determine which Hex or HexNAc was present.[9,14] A greater understanding of the dissociation chemistry of these analytes is one way to improve the accuracy of structural assignments and subsequent confidence in claims based on these assignments.[9,21,22]

We recently investigated the dissociation chemistry of single-residue glycopeptide model systems for the core structures of N-linked glycosylation, mucin-type glycosylation, and O-GlcNAcylation.[22] O-linked glycosylation is more varied than N-linked glycosylation due to the lack of either core glycan structures and the lack of consensus peptide sequences to ease identification of potential sites of glycosylation.[23] However, O-linked glycosylation to an alpha N-Acetyl galactose, α-GalNAc[17,24,25], residue is the most common residue and linkage type combination in humans and is also the subject of the present manuscript. Dissociation of these analytes produces GalNAc reporter ions at *m/z* 204 (Scheme 1).[23]



Scheme 1. Simplified formation of *m/z* 204 α-GalNAc oxazolinium reporter ion structures from protonated glycopeptides (R = peptide): (a) Direct, pyranose B_1 ion;[26–29] (b) Indirect, furanose B_1 ion[22].

Recently we provided mass spectral, isotopic (H/D) labelling, and computational evidence indicating that the dissociation chemistry of glycosylated amino acid model systems was more complex than had previously been believed.[22] Earlier labelling and tandem mass spectrometric work had proposed mechanisms for these dissociation processes.[27] In contrast, further, extensive ^{13}C and ^{15}N labelling and ion-mobility data from Mookherjee et al.[29] on bare N-Acetylated hexose model systems was not consis-

tent with these early mechanistic proposals.[27] i.e., HexNAc ions not attached to a peptide/amino acid. Our findings[22] were more consistent with the Mookherjee et al..[29] We found evidence for pyranose to furanose isomerization prior to glycosidic bond cleavage. i.e., that furanose oxazolinium B_1 ions are the predominant *m/z* 204 ion structure within the gas-phase populations (Scheme 1b). The newly released carbons 5 and 6 in the furanose structures are then available for secondary fragmentation to form the *m/z* 126 peak (loss of $C_2H_4O_2$ and H_2O) utilized to differentiate HexNAc-linked peptides based on MS^3 abundances.[22,26–28]

METHODS SECTION

Chemicals. HPLC grade Acetonitrile, HPLC grade water, and the formic acid were purchased from Sigma Aldrich (St. Louis, MO). Glycopeptides were custom synthesized by Karebay Biochemical, South Brunswick Township, New Jersey. All samples were used as received.

Tandem Mass Spectrometry. An electrospray ionization (ESI) Bruker MaXis plus quadrupole time-of-flight mass spectrometer (Bruker, Billerica, MA) was utilized for the initial MS and MS/MS experiments. Nitrogen was used as both nebulizing and drying gas in the ESI source. MS/MS spectra were obtained by quadrupole isolation of the precursor ion (E.g., $[AT(GalNAc)A+H]^+$, *m/z* 465.1±4) followed by collision-induced-dissociation (CID) with nitrogen in the hexapole collision cell. Product ion dispersion was achieved by the time-of-flight mass analyzer. Data were collected as a function of collision energy (averages of 120 spectra presented). Ionization was by electrospray with the samples infused into the instrument in ~1 μM acetonitrile/water/formic acid (50/50/0.1) solutions at a flow rate of 3 μl min^{-1}.

IRMPD Spectroscopy. Spectra were collected at the Centre Laser Infrarouge d'Orsay[30,31] in the X-H (X=N, O) stretching region (3200-3700 cm^{-1}) provided by an optical parametric oscillator/amplifier (OPO/OPA from LaserVision) laser system[32,33] light source. An Innolas Spitlight 600 non-seeded Nd:YAG (1064 nm, 550 mJ/pulse, bandwidth ~1 cm^{-1}) laser running at 25 Hz and delivering pulses of 4–6 ns duration pumps this system. Typical output energy of the OPO/OPA was 12–13 mJ/pulse at 3600 cm^{-1} with a 3–4 cm^{-1} (FWHM) bandwidth.

Experimental spectroscopic work was carried out in a Bruker Apex IV Qe, a 7 Tesla Fourier transform-ion cyclotron resonance (FT-ICR) tandem mass spectrometer (Bruker, Bremen, Germany)[32]. Precursor ions were pulse-extracted into the ICR cell where they were irradiated with IR light. With the OPO/OPA tuned on a vibrational transition of the mass-selected ion a significant boost in signal-to-noise is achieved by irradiating the ions for a few ms with a CO_2 laser pulse [10 Watt continuous wave (CW), BFi OPTiLAS, France] following each OPO/OPA pulse with a delay of ~1 μs. The total irradiation period was 1 s. This combination has been used previously.[34–38]

The abundances of the precursors and their corresponding photo fragments were recorded at each IR wavenumber, which was scanned stepwise. IR action spectra were derived by plotting the IRMPD efficiency against the photon energy. A Savitzky-Galoy[39] filter, with a rolling window length of 5 and a quadratic polynomial fit was used to smooth the raw datapoints.

IR action spectra were collected for the *m/z* 204.087 ($C_8H_{14}N_1O_5^+$), produced from $[AT(GalNAc)A+H]^+$, $[PT(GalNAc)P+H]^+$, and $[RVT(GalNAc)AG+2H]^{2+}$. i.e., the glycan fragment. The IR action spectrum of the *m/z* 262.143 $[ATA+H]^+$, $[AT(GalNAc)A+H\text{-}C_8H_{14}N_1O_5]^+$ peptide fragment was also collected. These ions were formed by low-energy collisional activation and thermalization with Ar (~10^{-3} mbar) in the linear hexapole. For the *m/z* 204.087 ions the IRMPD fragments followed were *m/z* 186.076, 168.066, 144.066, and 126.055. For $[ATA+H]^+$, the fragments followed were *m/z* 244.129 and 173.092.

Theoretical Methods. Similar to prior systems[22,40–44] we performed quantum chemical calculations to model the potential energy surfaces of $[AT(GalNAc)A+H]^+$, $[ATA+H]^+$, $[GalNAc\text{-}H_2O+H]^+$, ATA, and ($GalNAc\text{-}H_2O$). Initial candidate structures were systematically generated via the tool Fafoom[45–48], a genetic algorithm with the generated structures initially optimized using the MMFF94 Force Field[49–53]. Geometry optimizations of the resulting candidate conformations were performed with the Gaussian 09 software package[54] at four increasing levels of theory: HF/3-21g, B3LYP/6-31G(d), B3LYP/6-31+G(d,p),[55–57] and M06-2X/6-31+G(d,p)[58,59]. Degenerate structures were removed after each level and the remaining structures were utilized as the starting points of the subsequent calculation set. Supplemental manual manipulation and adjustment of the lowest energy structures followed by re-optimizations were also undertaken to ensure that other important conformers had not been missed.

Calculations of reaction pathways including transition structures, product ions and neutrals were performed at the B3LYP/6-31G(d), B3LYP/6-31+G(d,p), and M06-2X/6-31+G(d,p) levels of theory. Multiple transition structures (TSs) were systematically generated and calculated from multiple precursor ion structures for each potential fragmentation pathway. All minima and TSs were tested by vibrational analysis (all real frequencies or 1 imaginary frequency, respectively). The potential energy surface generated combines the zero-point energy correction (ZPE) to the electronic energy ($E_{el,0K}$) for improved accuracy ($\Delta E_{el+ZPE,0K}$). The related, standard enthalpy (ΔH_{298K}), Gibbs free energy (ΔG_{298K}), and entropy (ΔS_{298K}) corrections to 298 K were also determined. The reaction pathway through each TS was tested with intrinsic reaction coordinate calculations (GS2 keyword) with up to 10 steps in each direction. The terminating points of these calculations (one on product-side, one on reactant-side) were then optimized further to determine which minima were connected to each TS. Targeted single point calculations utilizing the B3LYP, M06-2X, and ωB97XD[54,60] models with 6-311+G(2d,2p) basis sets to assess the extent of energetic variability followed.

B3LYP/6-311+G(2d,p) vibrational frequencies from optimized structures at the same level of theory were utilized for comparisons with the experimental IRMPD spectra. A 20 cm^{-1} full width at half maximum Gaussian line shape and a scaling factor of 0.960 was utilized for the vibrational frequencies for comparison to the experimental spectra.

Description of Reactions. Roman numerals are utilized to label calculated structures (consistent with other ACS journals) in the text. Transition structures are labelled in a similar manner as: TS_reactant_product. For example: TS_I_II would be the TS structure linking reactant structure I to product structure II.

RESULTS AND DISCUSSION

Tandem Mass Spectrometry. For $[AT(GalNAc)A+H]^+$ our electrospray-quadrupole time-of-flight tandem mass spectra show abundant peaks at *m/z* 204 and 262 (**Figure 1**). These peaks are complementary from a single cleavage of the glycosidic bond: *m/z* 204 is $[GalNAc\text{-}H_2O+H]^+$, $C_8H_{14}N_1O_5^+$, and *m/z* 262 is protonated alanylthreonylalanine, $[ATA+H]^+$, $C_{10}H_{20}N_3O_5^+$. Facile cleavage of the glycosidic bond to produce substantial peaks at m/z 204 is consistent with the literature.[1,5,14,27,28,61] At increased collision energies consecutive losses from these ions become increasingly prevalent (m/z 186, 138, 126).

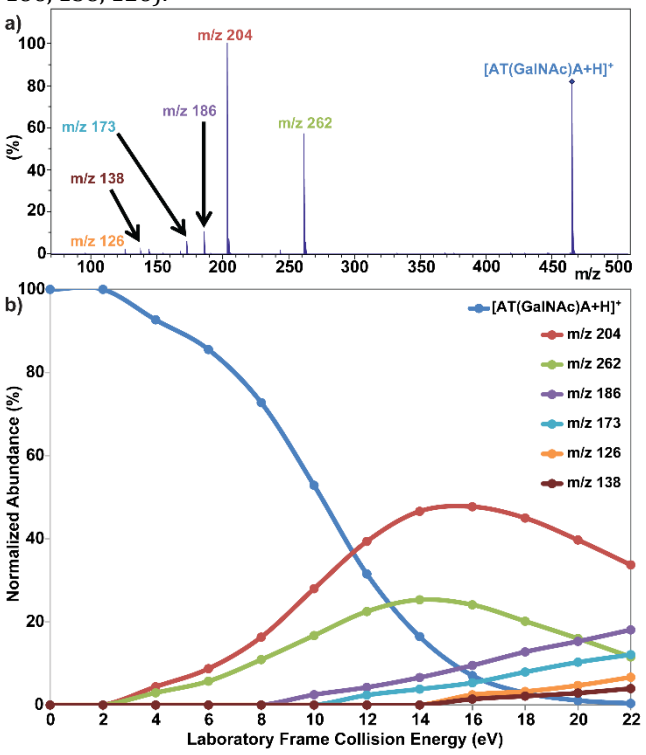


Figure 1. Electrospray-quadrupole time-of-flight MS/MS data: (a) Example spectrum of $[AT(GalNAc)A+H]^+$, *m/z* 465.1±4, at 12 eV laboratory frame collision energy. (b) Breakdown graph summarizing the $[AT(GalNAc)A+H]^+$ MS/MS spectra as a function of laboratory frame collision energy.

$[AT(GalNAc)A+H]^+$ Minima. The lowest energy conformer of $[AT(GalNAc)A+H]^+$ (Figure 2, Table S1) is predicted to have a highly folded structure[62] offering excellent charge solvation through hydrogen bonding side-chain N-acetyl and C-terminal carbonyl oxygens to the protonated N-terminus. These beneficial interactions result in a skew hexose ring which in turn has an additional hydrogen bonding network between the hydroxyls. Other low energy structures (Figure S1, Table S1) are predicted to be a mixture of highly folded (II) and more extended, so entropically favorable conformations (III, IV).

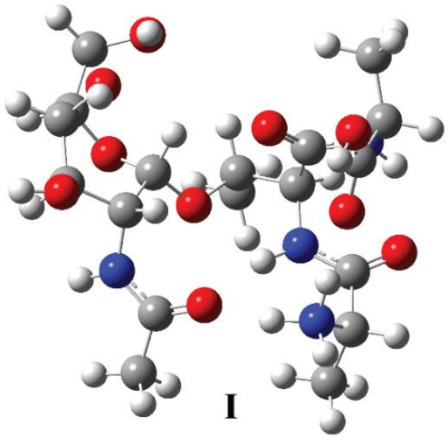
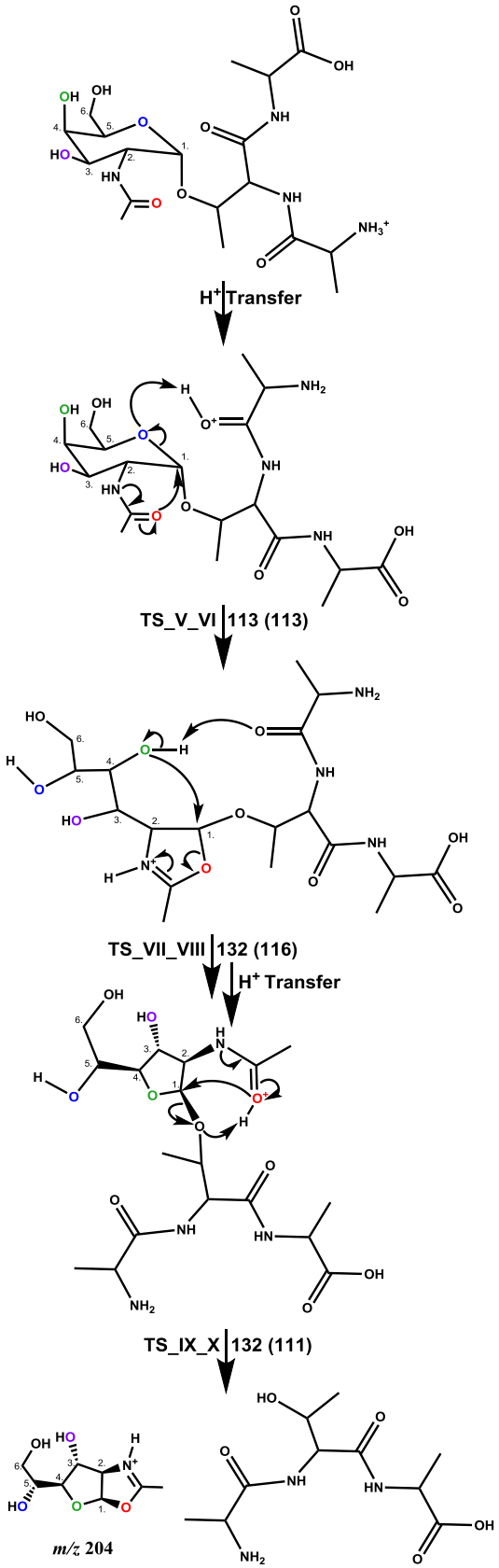


Figure 2. Global minimum precursor ion structure, I, of $[AT(GalNAc)A+H]^+$.

Glycosidic Bond Cleavage. The *m/z* 204, $C_8H_{14}N_1O_5^+$, $[GalNAc\text{-}H_2O+H]^+$ and *m/z* 262, $C_{10}H_{20}N_3O_5^+$, $[ATA+H]^+$ fragments are complementary resulting from cleavage of the glycosidic bond. Consistent with the wider understanding of peptide dissociation chemistry, the key difference is which fragment keeps the ionizing proton.[63–65]

For the $[AT(GalNAc)A+H]^+$ ions, the lowest energy mechanism involves isomerization[22,66] prior to glycosidic bond dissociation (**Scheme 2**). This pathway begins with mobilization of the ionizing proton to the carbonyl oxygen of the N-terminal alanine. The reactive configuration positions the pyranose ring oxygen solvating the ionizing proton (V, Figure S2, Table S1). Transfer of the proton to the ring oxygen and concerted nucleophilic attack into carbon 1 of the GalNAc residue in a S_N2-like transition structure (TS_V_VI, ΔH_{298K} (ΔG_{298K}) = 113 (113) kJ mol^{-1} (M06-2X/6-31+G(d,p)), Table S1, **Figure 3**) forms an oxazolinium-derivative that allows free rotation of carbons 3 to 6. Once the hydroxyl group of carbon 4 is appropriately positioned a second S_N2-like transition structure may occur (TS_VII_VIII, 131 (116) kJ mol^{-1}, Figure 3, Table S1) to produce a furanose GalNAc residue and protonation back at the carbonyl oxygen of the N-terminal alanine. Further proton transfers and rotations enable population of the furanose glycosidic bond cleavage reactive configuration (IX, Table S1, Figure S2). This elongated structure positions the ionizing proton close to the glycosidic oxygen while simultaneously bridged between the threonyl carbonyl oxygen and the carbonyl oxygen of the GalNAc residue. Proton transfer weakens the glycosidic bond,[21,22] enabling cleavage with concerted oxazolinium ion formation through nucleophilic attack by the carbonyl oxygen of GalNAc into carbon 1 of the GalNAc residue (TS_IX_X, 132 (111) kJ mol^{-1}, Table S1, Figure 3, Figure S2). The resulting proton bound dimer of ATA and the furanose oxazolinium ion, X, then either dissociates directly to yield the m/z 204 ion and neutral ATA or transfers a proton to ATA prior to dissociation, thereby generating m/z 262, $[ATA+H]^+$. For all these reactions the peptide provides key charge solvation to lower the barriers substantially.[22] Without this

solvation this complicated dissociation mechanism would not be feasible.



Scheme 2. The predicted lowest energy pathways to glycosidic bond dissociation *m/z* 204 α-GalNAc oxazolinium ion formation. The ΔH_{298K} (ΔG_{298K}) barriers are in kJ mol^{-1}.

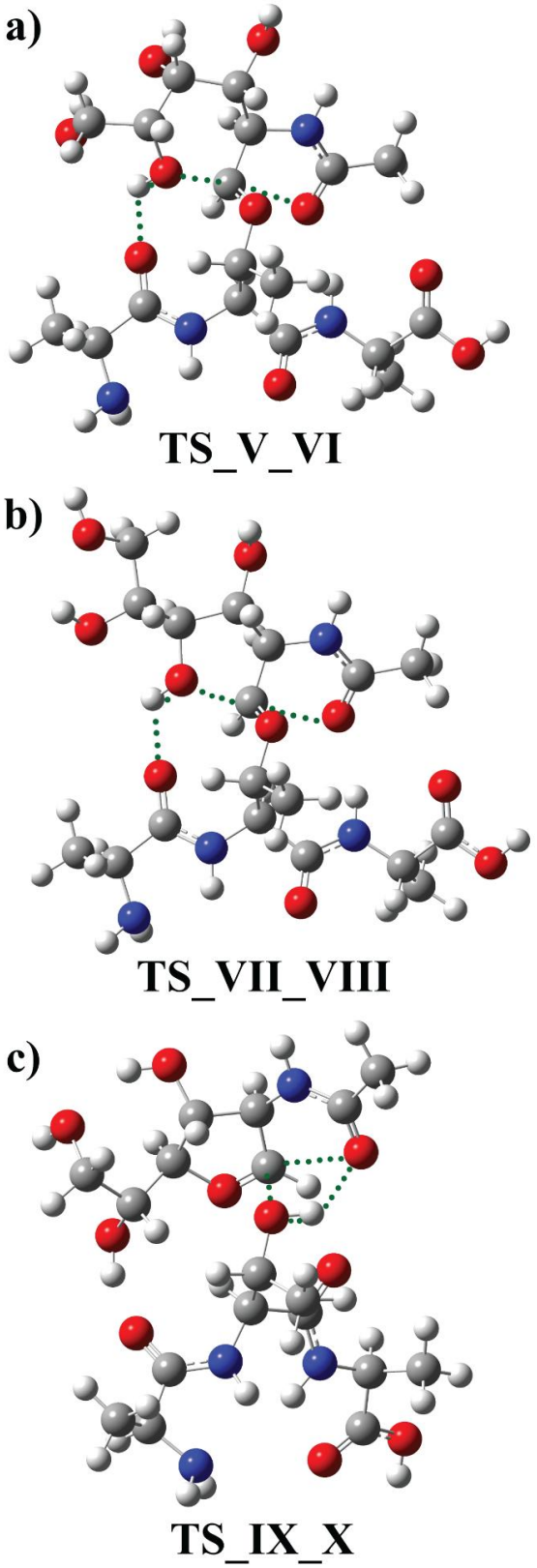


Figure 3. Transition structures enabling furanose, GalNAc oxazolinium ion, m/z 204 formation: (a) pyranose ring-opening; (b) S_N2-like formation of furanose GalNAc; (c) Proton transfer and glycosidic bond cleavage. Dotted green lines indicate bond breakage/formation.

The conceptually simpler direct mechanism[26–28] of dissociation (**Scheme S1**) is again[22] predicted to be more energetically demanding (TS_XI_XII, 143 (130) kJ mol^{-1}, Table S1, Figure S3). The mechanism begins with mobilization of a proton from the global minimum, N-terminally protonated structure to the glycosidic oxygen via the carbonyl oxygen of the GalNAc residue enabling concerted glycosidic bond cleavage and nucleophilic attack into carbon 1 to form a pyranose GalNAc oxazolinium ion. A summary of the 2 pathways and their relative energetics is provided in Figure S4.

The predicted reaction energetics are consistent for both the M06-2X and B3LYP models with 6-31+G(d,p) basis sets although the magnitude of the predicted barriers differs. To assess the effect of chemical model as a function of basis set, we utilized the larger 6-311+G(2d,2p) basis set and the ωB97XD functional too. All 3 models consistently predict the furanose-forming mechanism be most likely (Table S2). Like the Kuo group,[67] we find that the newer models which contain dispersion terms and larger proportions of exact exchange produce larger barriers for gas-phase glycan reactions.[21,22,41] Here, the key reactions are S_N2-like which is a known weakness of the B3LYP model which systematically underestimates S_N2 reaction barriers.[68–71]

IR Action Spectra of dissociation products: m/z 204. Figure 4 shows our experimental IR action spectroscopy data for the m/z 204 population compared to theoretical spectra of the six lowest energy furanose oxazolinium ion structures predicted within 20 kJ mol^{-1} (Figure S5). All theoretical spectra have at least one intense band predicted at ~3460 cm^{-1} corresponding to the NH stretch, and higher energy bands predicted in the 3640-3680 cm^{-1} region corresponding to free OH stretches. H-bonding pattern strongly affects which OH stretch this corresponds to. i.e., all three hydroxyl stretches are predicted as the highest energy band for at least one structure in Figure 4 (Table S3). No individual structure's theoretical spectrum explains all experimental features. Population of several low energy structures is more plausible. In particular, one low energy structure (panel d, ~3525 cm^{-1}, C5-OH hydrogen bonded to C6-OH, Table S3) predicts a substantial band assignable to the experimental feature ~3510-3550 cm^{-1}; conversely, the experimentally lower signal-to-noise band at ~3580 cm^{-1} is only predicted for one low energy structure (panel g, C5-OH hydrogen bonded to the free C6-OH, Table S3). What about the typically invoked[26–28] pyranose oxazolinium ion structures?

Calculated IR spectra of the lowest energy pyranose oxazolinium ion structures (Figure S6) are compared to the experimental spectrum in Figure S7. Both experimental and theoretical spectra show a substantial band at ~3460 cm^{-1} again corresponding to the NH stretch. However, no pyranose oxazolinium structure has a predicted band which could explain the intense experimental at ~3510-3550 cm^{-1}. The much less intense experimental feature at 3575-3600 cm^{-1} is potentially consistent with an intense band predicted for the lowest energy pyranose oxazolinium ion structure (Figure S7, panel f). Although we cannot rule out a population of the pyranose oxazolinium ion structures based on these data, these structures are insufficient to explain all experimental features.

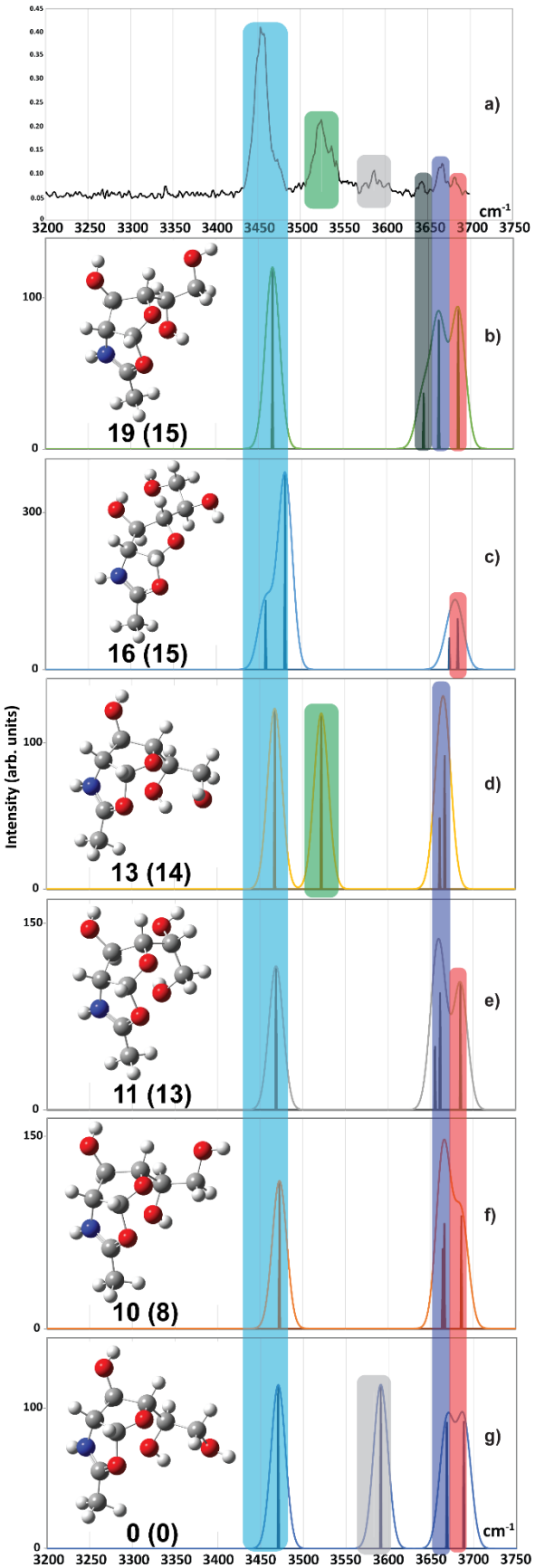


Figure 4. IR action spectra of the *m/z* 204 population (a) compared to the six lowest energy furanose GalNAc oxazolinium ion structures (b-g).

Additionally, data from the *m/z* 204 peaks produced from 2 other protonated glycopeptides, $[PT(GalNAc)P+H]^+$ and $[RVT(GalNAc)AG+2H]^{2+}$ collected on a subsequent days produced similar, but noisier spectra (Figure S8) to those in Figure 4. While these spectra are lower signal-to-noise, the consistency in band position supports a similar distribution of m/z 204 ion gas-phase ion populations despite changes in peptide sequence and charge state. i.e., indirect support for current MS^3-based approaches to residue identification.[26,27]

IR Action Spectra of dissociation products: m/z 262. Figure S9 shows our experimental IR action spectroscopy data for the m/z 262 population compared to the 6 lowest energy theoretical IR action spectra of $[ATA+H]^+$. These are O-protonated structures of similar energy (range <10 kJ mol^{-1}, Figure S10) and band position making definitively excluding structures difficult. The enthalpically lowest energy structure (Figure S9, panel f, Figure S10) is insufficient to explain the higher energy bands in the spectrum as it lacks a band predicted at ~3665 cm^{-1}. This is consistent with the greater entropic favorability of these structures (Figure S10).

Conclusions. Our combined tandem mass spectrometric, computational, and spectroscopic analyses provide evidence indicating that protonated α-GalNAc glycosylated peptides can produce furanose oxazolinium ion, *m/z* 204, $C_8H_{14}N_1O_5^+$, structures. Multiple levels of theory support isomerization prior to glycosidic bond dissociation to produce this ion. In principle, these reactions should be general for glycopeptides as the primary function of the peptide is to facilitate dissociation by providing charge solvation. However, multiple, low energy structures are necessary to explain the IR action spectrum. Concomitantly, we are unable to entirely rule out simultaneous production of at least some pyranose oxazolinium ion population due a combination of similar predicted bands and reaction barriers not being drastically different.

Future work will involve testing the generality of these findings for a wider range of glycan linkage types, glycan and peptide compositions, and at differing photon energies.

ASSOCIATED CONTENT

Supporting Information

Supplementary Tables summarizing relative energetics, Figures, minima, and transition structure figures, coupled with additional discussion of the dissociation energetics are provided in the supporting PDF file. The Supporting Information is available free of charge on the ACS Publications website.

AUTHOR INFORMATION

Corresponding Author

*Benjamin Bythell, bythell@ohio.edu

Author Contributions

[Δ]JMR and SG contributed equally to this manuscript. JMR, SG, and BJB performed the calculations. MTA and SG collected and initially analyzed the tandem mass spectral data. JMR, BJB, and PM collected the spectroscopic data. All authors analyzed data. Figures and Schemes were made by LMS, JMR, and BJB. The manuscript was written through contributions of all authors.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under CHE-1948611, a University of Missouri-St. Louis Research Award, and start-up funds from Ohio University. Calculations were performed at Ohio Supercomputer Center (http://osc.edu/ark:/19495/f5s1ph73) and at the Missouri University of Science and Technology, Rolla, MO with support from the NSF (ACI-1919789). The research leading to this result has been supported by the project CALIPSOplus under the Grant Agreement 730872 from the EU Framework Program for Research and Innovation HORIZON 2020. The authors are grateful to E. Loire, J. M. Ortega, F. Gobert, and N. Jestin for technical support.

REFERENCES

(1) Bones, J.; Mittermayr, S.; O'Donoghue, N.; Guttman, A.; Rudd, P. M. Ultra Performance Liquid Chromatographic Profiling of Serum N-Glycans for Fast and Efficient Identification of Cancer Associated Alterations in Glycosylation. *Anal. Chem.* **2010**, *82* (24), 10208–10215. https://doi.org/10.1021/ac102860w.

(2) Saldova, R.; Royle, L.; Radcliffe, C.; Abd hamid, U.; Evans, R.; Arnold, J.; Banks, R.; Huston, R.; Harvey, D.; Antrobus, R.; Petrescu, S.; Dwek, R.; Rudd, P. Ovarian Cancer Is Associated with Changes in Glycosylation in Both Acute-Phase Proteins and IgG. *Glycobiology* **2007**, *17* (12), 1344–1356.

(3) Miyoshi, E.; Nakano, M. Fucosylated Haptoglobin Is a Novel Marker for Pancreatic Cancer: Detailed Analyses of Oligosaccharide Structures. *Proteomics* **2008**, *8* (16), 3257–3262. https://doi.org/10.1002/pmic.200800046.

(4) Okuyama, N.; Ide, Y.; Nakano, M.; Nakagawa, T.; Yamanaka, K.; Moriwaki, K.; Murata, K.; Ohigashi, H.; Yokoyama, S.; Eguchi, H.; Ishikawa, O.; Ito, T.; Kato, M.; Kasahara, A.; Kawano, S.; Gu, J.; Taniguchi, N.; Miyoshi, E. Fucosylated Haptoglobin Is a Novel Marker for Pancreatic Cancer: A Detailed Analysis of the Oligosaccharide Structure and a Possible Mechanism for Fucosylation. *Int. J. Can-*

cer **2006**, *118* (11), 2803–2808. https://doi.org/10.1002/ijc.21728.

(5) Dube, D. H.; Bertozzi, C. R. Glycans in Cancer and Inflammation - Potential for Therapeutics and Diagnostics. *Nat. Rev. Drug Discov.* **2005**, *4* (6), 477–488. https://doi.org/10.1038/nrd1751.

(6) Rachagani, S.; Torres, M. P.; Moniaux, N.; Batra, S. K. Current Status of Mucins in the Diagnosis and Therapy of Cancer. *BioFactors* **2009**, *35* (6), 509–527. https://doi.org/10.1002/biof.64.

(7) Cazet, A.; Julien, S.; Bobowski, M.; Krzewinski-Recchi, M. A.; Harduin-Lepers, A.; Groux-Degroote, S.; Delannoy, P. Consequences of the Expression of Sialylated Antigens in Breast Cancer. *Carbohydr. Res.* **2010**, *345* (10), 1377–1383. https://doi.org/10.1016/j.carres.2010.01.024.

(8) Tian, Y.; Zhang, H. Characterization of Disease-Associated N -Linked Glycoproteins. *PROTEOMICS* **2012**, *13* (3–4), 504–511. https://doi.org/10.1002/pmic.201200333.

(9) Chernykh, A.; Kawahara, R.; Thaysen-Andersen, M. Towards Structure-Focused Glycoproteomics. *Biochem. Soc. Trans.* **2021**, *49* (1), 161–186. https://doi.org/10.1042/BST20200222.

(10) Gornik, O.; Royle, L.; Harvey, D.; Radcliffe, C.; Saldova, R.; Dwek, R.; Rudd, P.; Lauc, G. Changes of Serum Glycans during Sepsis and Acute Pancreatitis. *Glycobiology* **2007**, *17* (12), 1321–1332.

(11) Sperandio, M.; Gleissner, C. A.; Ley, K. Glycosylation in Immune Cell Trafficking. *Immunol. Rev.* **2009**, *230* (1), 97–113. https://doi.org/10.1111/j.1600-065X.2009.00795.x.

(12) Shajahan, A.; Pepi, L. E.; Rouhani, D. S.; Heiss, C.; Azadi, P. Glycosylation of SARS-CoV-2: Structural and Functional Insights. *Anal. Bioanal. Chem.* **2021**, *413* (29), 7179–7193. https://doi.org/10.1007/s00216-021-03499-x.

(13) Arnaud, C. H. Adding the Missing Sugars to Coronavirus Protein Structures. *Chem. Eng. News* **2021**, *98* (16).

(14) Darula, Z.; Medzihradszky, K. F. Analysis of Mammalian O-Glycopeptides - We Have Made a Good Start, but There Is a Long Way to Go. *Mol. Cell. Proteomics* **2018**, *17* (1), 2–17. https://doi.org/10.1074/mcp.MR117.000126.

(15) You, X.; Qin, H.; Ye, M. Recent Advances in Methods for the Analysis of Protein O-Glycosylation at Proteome Level. *J. Sep. Sci.* **2018**, *41* (1), 248–261. https://doi.org/10.1002/jssc.201700834.

(16) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, *126* (5), 855–867. https://doi.org/10.1016/j.cell.2006.08.019.

(17) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in Health and Disease. *Nat. Rev. Nephrol.* **2019**, *15* (6), 346–366. https://doi.org/10.1038/s41581-019-0129-4.

(18) Peltoniemi, H.; Joenväärä, S.; Renkonen, R. De Novo Glycan Structure Search with the CID MS/MS Spectra of Native N-Glycopeptides. *Glycobiology* **2009**, *19* (7), 707–714. https://doi.org/10.1093/glycob/cwp034.

(19) Ozohanics, O.; Krenyacz, J.; Ludányi, K.; Pollreisz, F.; Vékey, K.; Drahos, L. GlycoMiner: A New Software Tool to Elucidate Glycopeptide Composition. *Rapid Commun. Mass Spectrom.* **2008**, *22* (20), 3245–3254. https://doi.org/10.1002/rcm.3731.

(20) Goldberg, D.; Bern, M.; Parry, S.; Sutton-Smith, M.; Panico, M.; Morris, H. R.; Dell, A. Automated N-Glycopeptide Identification Using a Combination of Single- and Tandem-MS. *J. Proteome Res.* **2007**, *6* (10), 3995–4005. https://doi.org/10.1021/pr070239f.

(21) Bythell, B. J.; Abutokaikah, M. T.; Wagoner, A. R.; Guan, S.; Rabus, J. M. Cationized Carbohydrate Gas-Phase Fragmentation Chemistry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 688–703.

(22) Guan, S.; Bythell, B. J. Evidence of Gas-Phase Pyranose-to-Furanose Isomerization in Protonated Peptidoglycans. *Phys. Chem. Chem. Phys.* **2021**, *23*, 23256–23266. https://doi.org/10.1039/D1CP03842G.

(23) Zhu, Z.; Desaire, H. Carbohydrates on Proteins: Site-Specific Glycosylation Analysis by Mass Spectrometry. *Annu. Rev. Anal. Chem.* **2015**, *8* (1), 463–483. https://doi.org/10.1146/annurev-anchem-071114-040240.

(24) Jensen, P. H.; Kolarich, D.; Packer, N. H. Mucin-Type O-Glycosylation - Putting the Pieces Together. *FEBS J.* **2010**, *277* (1), 81–94. https://doi.org/10.1111/j.1742-4658.2009.07429.x.

(25) Alley, W. R.; Novotny, M. V. Structural Glycomic Analyses at High Sensitivity: A Decade of Progress. *Annu. Rev. Anal. Chem.* **2013**, *6* (1), 237–265. https://doi.org/10.1146/annurev-anchem-062012-092609.

(26) Halim, A.; Westerlind, U.; Pett, C.; Schorlemer, M.; Rüetschi, U.; Brinkmalm, G.; Sihlbom, C.; Lengqvist, J.; Larson, G.; Nilsson, J. Assignment of Saccharide Identities through Analysis of Oxonium Ion Fragmentation Profiles in LC–MS/MS of Glycopeptides. *J. Proteome Res.* **2014**, *13* (12), 6024–6032. https://doi.org/10.1021/pr500898r.

(27) Yu, J.; Schorlemer, M.; Gomez Toledo, A.; Pett, C.; Sihlbom, C.; Larson, G.; Westerlind, U.; Nilsson, J. Distinctive MS/MS Fragmentation Pathways of Glycopeptide-Generated Oxonium Ions Provide Evidence of the Glycan Structure. *Chem. – Eur. J.* **2016**, *22* (3), 1114–1124. https://doi.org/10.1002/chem.201503659.

(28) Ashwood, C.; Lin, C.-H.; Thaysen-Andersen, M.; Packer, N. H. Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. *J. Am. Soc. Mass Spectrom.* **2018**, 1–16. https://doi.org/10.1007/s13361-018-1932-z.

(29) Mookherjee, A.; Uppal, S. S.; Guttman, M. Dissection of Fragmentation Pathways in Protonated *N* -Acetylhexosamines. *Anal. Chem.* **2018**, *90* (20), 11883–11891. https://doi.org/10.1021/acs.analchem.8b01963.

(30) Prazeres, R.; Glotin, F.; Insa, C.; Jaroszynski, D. A.; Ortega, J. M. Two-Colour Operation of a Free-Electron Laser and Applications in the Mid-Infrared. *Eur. Phys. J. - At. Mol. Opt. Plasma Phys.* **1998**, *3* (1), 87–93. https://doi.org/10.1007/s100530050151.

(31) Lemaire, J.; Boissel, P.; Heninger, M.; Mauclaire, G.; Bellec, G.; Mestdagh, H.; Simon, A.; Caer, S. L.; Ortega, J. M.; Glotin, F.; Maitre, P. Gas Phase Infrared Spectroscopy of Selectively Prepared Ions. *Phys. Rev. Lett.* **2002**, *89* (27), 273002. https://doi.org/10.1103/PhysRevLett.89.273002.

(32) Bakker, J. M.; Besson, T.; Lemaire, J.; Scuderi, D.; Maître, P. Gas-Phase Structure of a π-Allyl−Palladium Complex: Efficient Infrared Spectroscopy in a 7 T Fourier Transform Mass Spectrometer. *J. Phys. Chem. A* **2007**, *111* (51), 13415–13424. https://doi.org/10.1021/jp074935e.

(33) Bakker, J. M.; Sinha, R. K.; Besson, T.; Brugnara, M.; Tosi, P.; Salpin, J.-Y.; Maître, P. Tautomerism of Uracil Probed via Infrared Spectroscopy of Singly Hydrated Protonated Uracil. *J. Phys. Chem. A* **2008**, *112* (48), 12393–12400. https://doi.org/10.1021/jp806396t.

(34) Yeh, L.I.; Okumura, M.; Myers, J.D.; Price, J.M.; Lee, Y.T. Vibrational Spectroscopy of the Hydrated Hydronium Cluster Ions H3O+·(H2O)n (N=1, 2, 3): *J. Chem. Phys.* **1989**, *91* (12), 7319–7330.

(35) Sinha, R. K.; Nicol, E.; Steinmetz, V.; Maître, P. Gas Phase Structure of Micro-Hydrated [Mn(ClO4)]+ and [Mn2(ClO4)3]+ Ions Probed by Infrared Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (5), 758–772. https://doi.org/10.1016/j.jasms.2010.02.014.

(36) Sinha, R. K.; Erlekam, U.; Bythell, B. J.; Paizs, B.; Maître, P. Diagnosing the Protonation Site of $\boldsymbol{b}_2$ Peptide Fragment Ions Using IRMPD in the X–H (X = O, N, and C) Stretching Region. *J. Am. Soc. Mass*

Spectrom. **2011**, *22* (9), s13361-011-0173–1. https://doi.org/10.1007/s13361-011-0173-1.

(37) Peiris, D. M.; Cheeseman, M. A.; Ramanathan, R.; Eyler, J. R. Infrared Multiple Photon Dissociation Spectra of Gaseous Ions. *J Phys Chem* **1993**, *97* (30), 7839–7843. https://doi.org/10.1021/j100132a009.

(38) Lagutschenkov, A.; Sinha, R. K.; Maitre, P.; Dopfer, O. Structure and Infrared Spectrum of the Ag+─Phenol Ionic Complex†. *J Phys Chem A* **2010**, *114* (42), 11053–11059. https://doi.org/10.1021/jp100853m.

(39) Savitzky, Abraham.; Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36* (8), 1627–1639. https://doi.org/10.1021/ac60214a047.

(40) Abutokaikah, M. T.; Frye, J. W.; Tschampel, J.; Rabus, J. R.; Bythell, B. J. Fragmentation Pathways of Lithiated Hexose Monosaccharides. *J Am Soc Mass Spectrom* **2018**, *29*, 1627–1637.

(41) Bythell, B. J.; Rabus, J. M.; Wagoner, A. R.; Abutokaikah, M. T.; Maitre, P. Sequence Ion Structures and Dissociation Chemistry of Deprotonated Sucrose Anions. *J Am Soc Mass Spectrom* **2018**, *29*, 2380–2393. https://doi.org/DOI: 10.1007/s13361-018-2065-0.

(42) Rabus, J. M.; Abutokaikah, M. T.; Ross, R. T.; Bythell, B. J. Sodium-Cationized Carbohydrate Gas-Phase Fragmentation Chemistry: Influence of Glycosidic Linkage Position. *Phys Chem Chem Phys* **2017**, *19* (37), 25643–25652. https://doi.org/10.1039/C7CP04738J.

(43) Rabus, J. M.; Simmons, D. R.; Maître, P.; Bythell, B. J. Deprotonated Carbohydrate Anion Fragmentation Chemistry: Structural Evidence from Tandem Mass Spectrometry, Infra-Red Spectroscopy, and Theory. *Phys. Chem. Chem. Phys.* **2018**, *20* (44), 27897–27909. https://doi.org/10.1039/C8CP02620C.

(44) Rabus, J. M.; Pellegrinelli, R. P.; Khodr, A. H. A.; Bythell, B. J.; Rizzo, T. R.; Carrascosa, E. Unravelling the Structures of Sodiated β-Cyclodextrin and Its Fragments. *Phys. Chem. Chem. Phys.* **2021**, *23* (24), 13714–13723. https://doi.org/10.1039/D1CP01058A.

(45) Supady, A. *adrianasupady/fafoom*. GitHub. https://github.com/adrianasupady/fafoom (accessed 2017-08-25).

(46) Marianski, M.; Supady, A.; Ingram, T.; Schneider, M.; Baldauf, C. Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α-Maltose. *J Chem Theory Comput* **2016**, *12*, 6157–6168.

(47) Supady, A.; Blum, V.; Baldauf, C. First-Principles Molecular Structure Search with a Genetic Algorithm. *J. Chem. Inf. Model.* **2015**, *55* (11), 2338–2348. https://doi.org/10.1021/acs.jcim.5b00243.

(48) Landrum, G.; Kelley, B.; Tosco, P.; sriniker; gedeck; NadineSchneider; Vianello, R.; Dalke, A.; Cole, B.; AlexanderSavelyev; Turk, S.; Swain, M.; Ric; Vaucher, A.; Wójcikowski, M.; N, D.; Pahl, A.; JP; strets123; JLVarjo; O'Boyle, N.; Berenger, F.; Fuller, P.; Jensen, J. H.; Sforna, G.; DoliathGavid; Cosgrove, D.; Nowotka, M.; Leswing, K.; Santen, J. van. Rdkit/Rdkit: 2018_09_3 (Q3 2018) Release, 2019. https://doi.org/10.5281/zenodo.2608859.

(49) Halgren, T. A. Merck Molecular Force Field. II. MMFF94 van Der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comput. Chem.* **1996**, *17* (5–6), 520–552. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<520::AID-JCC2>3.0.CO;2-W.

(50) Halgren, T. A. Merck Molecular Force Field. III. Molecular Geometries and Vibrational Frequencies for MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 553–586. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<553::AID-JCC3>3.0.CO;2-T.

(51) Halgren, T. A. Merck Molecular Force Field. V. Extension of MMFF94 Using Experimental Data, Additional Computational Data, and Empirical Rules. *J.*

Comput. Chem. **1996**, *17* (5–6), 616–641. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<616::AID-JCC5>3.0.CO;2-X.

(52) Halgren, T. A.; Nachbar, R. B. Merck Molecular Force Field. IV. Conformational Energies and Geometries for MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 587–615. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<587::AID-JCC4>3.0.CO;2-Q.

(53) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5–6), 490–519. https://doi.org/10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

(54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision E.01*; Gaussian, Inc.: Wallingford CT, 2009.

(55) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation Energy Formula into a Functional of the Electron Density. *Phys Rev B* **1988**, *37*, 785–789.

(56) Stephens, P. J.; Devlin, J. F.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J Phys Chem* **1994**, *98*, 11623–11627.

(57) Becke, A. D. Density- functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.

(58) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. Exchange-Correlation Functional with Broad Accuracy for Metallic and Nonmetallic Compounds, Kinetics, and Noncovalent Interactions. *J Chem Phys* **2005**, *123 (16)*, 161103.

(59) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor Chem Acc.* **2006**, *120*, 215.

(60) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–1799. https://doi.org/10.1002/jcc.20495.

(61) Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. *Nat. Chem. Biol.* **2010**, *6* (10), 713–723. https://doi.org/10.1038/nchembio.437.

(62) Mookherjee, A.; Van Stipdonk, M. J.; Armentrout, P. B. Thermodynamics and Reaction Mechanisms of Decomposition of the Simplest Protonated Tripeptide, Triglycine: A Guided Ion Beam and Computational Study. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 739–757. https://doi.org/10.1007/s13361-016-1590-y.

(63) Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and localized protons: a framework for understanding pep-

tide dissociation. *J Mass Spectrom* **2000**, *35*, 1399–1406.
(64) Tsaprailis, G.; Nair, H.; Somogyi, Á.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. Influence of Secondary Structure on the Fragmentation of Protonated Peptides. *J. Am. Chem. Soc.* **1999**, *121* (22), 5142–5154. https://doi.org/10.1021/ja982980h.
(65) Paizs, B.; Suhai, S. Fragmentation pathways of protonated peptides. *Mass Spectrom Rev* **2005**, *24*, 508–548.
(66) Abutokaikah, M. T.; Gnawali, G. R.; Frye, J. W.; Stump, C. M.; Tschampel, J.; Murphy, M. J.; Lachance, E. S.; Guan, S.; Spilling, C. D.; Bythell, B. J. Leaving Group Effects in a Series of Electrosprayed CcHhN1 Anthracene Derivatives. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2306–2317. https://doi.org/10.1007/s13361-019-02298-0.
(67) Chiu, C.; Lin, C.-K.; Kuo, J.-L. Improved Agreement between Experimental and Computational Results for Collision-Induced Dissociation Mass Spectrometry of Cation-Tagged Hexoses. *Phys. Chem. Chem. Phys.* **2020**, *22* (13), 6928–6941. https://doi.org/10.1039/D0CP00286K.
(68) Laerdahl, J. K.; Uggerud, E. Gas phase nucleophilic substitution. *Int J Mass SPectrom* **2002**, *214*, 277–314.
(69) Bythell, B. J.; Harrison, A. G. Formation of a_1 Ions Directly from Oxazolone b_2 Ions: An Energy-Resolved and Computational Study. *J Am Soc Mass Spectrom* **2015**, *26*, 774–781.
(70) Gritsenko, O. V.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J. Comparison of the Accurate Kohn-Sham Solution with the Generalized Gradient Approximations (GGAs) for the SN2 Reaction F- + CH3F -> FCH3 + F-: A Qualitative Rule To Predict Success or Failure of GGAs. *J Phys Chem A* **2000**, *104*, 8558–8565.
(71) Bleiholder, C.; Paizs, B. Competing Gas-Phase Fragmentation Pathways of Asparagine-, Glutamine-, and Lysine-Containing Protonated Dipeptides. *Theor Chem Acc* **2010**, *125*, 387–396.

TOC Figure: